DOCUMENT RESUME

ED 284 182 CS 008 873

AUTHOR Anderson, Richard C.; And Others
TITLE Do Errors on Classroom Reading Tasks Slow Growth in Reading? Technical Report No. 404.
INSTITUTION Bolt, Beranek and Newman, Inc., Cambridge, Mass.; Illinois Univ., Urbana. Center for the Study of Reading.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE Jan 87
CONTRACT 400-81-0030
NOTE 21p.; A version of this report will appear in the "Elementary School Journal," January 1988.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Elementary Education; *Error Analysis (Language); Miscue Analysis; *Oral Reading; Reading Achievement; Reading Comprehension; *Reading Processes; *Reading Research

ABSTRACT

A pervasive finding from research on teaching and classroom learning is that a low rate of error on classroom tasks is associated with large year to year gains in achievement, particularly for reading in the primary grades. The finding of a negative relationship between error rate, especially rate of oral reading errors, and gains in reading achievement is generally interpreted to mean that a low rate of error is on the causal path to growth in reading. However, this relationship may be an epiphenomenon: (1) error rate on classroom tasks is a good measure of children's level of reading development, (2) standardized tests and other one-shot assessment instruments are always imperfect measures of reading level, and (3) error rate correlates negatively with end-of-year achievement because it provides additional and more reliable information, beyond that contained in previous test scores, about children's reading ability. Findings from a microanalytic study of third-grade reading lessons confirmed that oral reading errors can have a positive influence on children's comprehension. Errors facilitated comprehension of nonturntakers when task norms emphasized accurate oral reading, but not when norms emphasized story understanding. According to the positive influence theory, an oral reading error followed by feedback fits the pattern of tension followed by resolution shown by other research to improve learning and memory. When the task is accurate oral reading, a failure to read fluently produces tension, which increases attention and instigates deeper processing or a greater "effort after meaning." (A table, figure and references are included.) (Author/HTH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED284182

CENTER FOR THE STUDY OF READING

Technical Report No. 404

DO ERRORS ON CLASSROOM READING TASKS SLOW GROWTH IN READING?

Richard C. Anderson, Ian Wilkinson, Jana M. Mason
University of Illinois at Urbana-Champaign
Larry Shirey
Santa Clara County Office of Education
Paul T. Wilson
Western Michigan University

January 1987

University of Illinois
at Urbana-Champaign
51 Gerty Drive
Champaign, Illinois 61820

Bolt Beranek and Newman Inc.
10 Moulton Street
Cambridge, Massachusetts 02238

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

The authors are indebted to Jere E. Brophy, Irene W. Gaskins, and Elfrieda H. Hiebert for helpful comments on earlier drafts of this paper. This work was supported in part under Contract No. 400-81-0030 from the National Institute of Education. A version of this paper is to appear in the Elementary School Journal, January 1988.

CS008873

Abstract

The purpose of this paper is a close examination of the evidence on error rate on classroom tasks. The finding of a negative relationship between error rate, especially rate of oral reading errors, and gains in reading achievement is generally interpreted to mean that a low rate of error is on the causal path to growth in reading. However, we argue that this relationship may be an epiphenomenon: (1) error rate on classroom tasks is a good measure of children's level of reading development, (2) standardized tests and other one-shot assessment instruments are always imperfect measures of reading level, (3) error rate correlates negatively with end-of-year achievement because it provides additional and more reliable information, beyond that contained in previous test scores, about children's reading ability. Findings from a microanalytic study of third-grade reading lessons confirmed that oral reading errors can have a *positive* influence on children's comprehension. Errors facilitated comprehension of nonturntakers when task norms emphasized accurate oral reading, but not when norms emphasized story understanding. According to our theory, an oral reading error followed by feedback fits the pattern of tension followed by resolution shown by other research to improve learning and memory. When the task is accurate oral reading, a failure to read fluently produces tension, which increases attention and instigates deeper processing or a greater "effort after meaning."

Do Errors on Classroom Reading Tasks Slow Growth in Reading?

A pervasive finding from research on teaching and classroom learning is that a low rate of error on classroom tasks is associated with large year to year gains in achievement (Rosenshine & Stevens, 1984). This finding holds particularly for reading in the primary grades. It generally is interpreted to mean that a low error rate is a *cause* of rapid progress in reading (Cunningham, 1985). Naturally, the conclusion that success causes growth in reading licenses several kinds of instructional advice, notably that children should be given easy books (Betts, 1946; Cunningham, 1985; Hoffman, O'Neal, Kastler, Clements, Segel, & Nash, 1984).

The purpose of this paper is a close examination of the evidence on error rate. We will argue that the well-known studies showing a relationship--maybe a causal relationship--between error rate and achievement can be given a plausible alternative interpretation. We will present evidence from recent studies involving *microanalysis* of classroom reading instruction that suggests that, far from impeding growth, an error during a reading lesson is an occasion for learning. A microanalysis, as we will show later, gives a moment-to-moment description of student behaviors that are intimately connected to the instructional environment and that are related to learning.

We begin with an overview of the major classroom studies that have examined the rate of student errors. *Error rate* is defined as the number of student responses judged incorrect or unsatisfactory divided by the total number of student responses recorded during a period of observation. Long hesitations or failures to respond are also negatively related to achievement gains and findings are comparable, whether or not these are included in error rate.

In the Beginning Teacher Evaluation Study, Fisher and his associates (Fisher, Filby, Marliave, Cahen, Dishaw, Moore, & Berliner, 1978) found that achievement gains were positively correlated with the percentage of class time during which the error rate was 10% or less and negatively correlated with the percentage of class time during which the error rate was greater than 30%. These findings held in both mathematics and reading across 25 second-grade classrooms and 21 fifth-grade classrooms. A high error rate had the greatest adverse effect in second-grade reading.

In the First-grade Reading Group Study, Anderson, Evertson, and Brophy (1979) found a strong positive relationship between the percentage of all correct student responses during reading lessons in 20, first-grade classrooms and the average end-of-year scores of these classrooms on a standardized reading achievement test, after controlling for scores on a test of reading readiness. They also found a negative relationship between oral reading errors and achievement, again after controlling for reading readiness. Among the many interesting and subtle results of this study was the finding that growth in reading is positively related to the teacher's providing *sustaining feedback*--furnishing hints to help the student come up with a satisfactory response, but negatively related to the teacher's giving *terminal feedback*--telling the student the correct response or calling on another student.

In a study of the high and low reading groups in 22 second-grade classrooms that investigated in great detail types of oral reading errors, the nature of teachers' feedback, and children's reactions following errors and feedback, Hoffman et al. (1984) confirmed that error rate and terminal feedback are negatively related to gains in reading achievement. The effects were smaller than in the Anderson et al. (1979) study, however. The reason may be that, because the error rates were low, the measure was skewed and had a curvilinear relationship with achievement. If so, the effects would have been stronger if error rates had been transformed. Hoffman et al. (1984) said they didn't transform the data because they didn't want to make an arbitrary decision about what value to assign to a zero error rate. Anderson et al. (1979) also observed a low rate of oral reading errors, and they didn't transform the measure either, though apparently they checked relationships for curvilinearity and none was reported in this instance.

To summarize, the major studies of classroom reading instruction that have included a measure of error rate have consistently found it to be negatively related to achievement gains. This finding is robust over variations in the design, procedure, and methods of analysis employed in these studies.

An Hypothesis About Error Rate and Reading Achievement

It is usually taken for granted that a low error rate is on the causal path to growth in reading. That this is so is evidenced by the fact that the presumed need for a low error rate is simply assumed in arguments about the desirability of clear directions, small steps, adequate amounts of practice, silent reading before oral reading, avoiding frustration and increasing feelings of self-efficacy, and, especially, the importance of easy reading materials. For instance, on the basis of the data on error rate and achievement, Hoffman et al. (1984) suggest that the long-time consensus in reading about the *instructional level* of materials--according to which materials are suitable for use in teacher-led instruction when students can orally read 95% of the words correctly--should be revised upward.

A deeper analysis shows that it *is* problematic that errors cause slower growth in reading. The first fact to consider is that covering more material and more difficult material is also associated with growth in reading (Anderson et al., 1979; Barr, Dreeben, & Wiratchai, 1983; Juel, Griffith, & Gough, 1986). Now, it is certainly true that children reading more difficult material will make more mistakes (Blaxall & Willows, 1984). Thus, on the face of things at least, a low error rate and a fast pace through material are inconsistent with one another. Considering this fact, the Commission on Reading felt obliged to caution in *Becoming a Nation of Readers* (Anderson, Hiebert, Scott, & Wilkinson, 1985) that ". . . How the most effective teachers manage to maintain both a fast pace and a high rate of success, two characteristics that may sometimes conflict, is a complex issue that research has not yet completely untangled" (p. 88).

It is possible that skillful teachers are able both to keep error rates low and to move rapidly through the curriculum--by not wasting time, maintaining high levels of attention, presenting informative explanations, and so on. However, there is another interpretation that we regard as highly plausible: (1) Error rate on daily classroom tasks is a good measure of children's level of reading development. (2) Standardized tests and other one-shot assessment instruments are always imperfect measures of reading level. (3) Error rate on classroom tasks correlates negatively with end-of-year achievement because it provides further information, beyond that contained in previous test scores, about children's level of reading ability.

Investigators who believe that they have "controlled" for initial reading level when they simply factor out scores on a pretest [through regression analysis, analysis of covariance, residual gain score analysis, or partial correlation analysis] make the mistake, whether they are aware of it or not, of assuming that their pretest is an infallible measure of initial level. This can be a serious mistake, and it is one that has been made by even sophisticated instructional researchers. For instance, we have shown that the otherwise worthy study by Leinhardt, Zigmond, and Cooley (1981) is marred by a data analysis that failed to take account of measurement error. Depending on the reliability of the pretest, time spent in silent reading could be shown to have a significant positive effect, no significant effect, or even a significant negative effect on students' posttest performance (Wilkinson, Wardrop, & Anderson, in press). Thus, quite different conclusions are reached when measurement error is systematically considered (see also Pedhazur, 1982).

It is well known that the younger the child the more fallible are tests of abilities (Butler, Marsh, Sheppard, & Sheppard, 1985; Walberg & Tsai, 1984). Thus, the explanation we are proposing readily accounts for the apparent fact that the negative correlation between error rate and achievement is largest in the first grade and is smaller in the second and higher grades. Otherwise, explanations are circular, no more than repetitions of the facts themselves: "The achievement of older students is less affected by error rate."

Reading readiness tests, in particular, are notoriously unreliable. This means that it is the first-grade studies, which have produced the most striking results of all instructional studies--such as those of Anderson et al. (1979), Barr et al. (1983), and Weinstein (1976)--that are most in doubt and most subject to reinterpretation. There are standard methods for making adjustments for error of measurement (see Linn & Werts, 1982; Zeller & Carmines, 1980). These should be used, where possible, in additional analyses of valuable data that already have been collected. It goes without saying that future investigations should pay attention to careful measurement, should never neglect to obtain good estimates of reliability, and where possible should obtain multiple indicators of reading level.

The task of stable measurement of initial reading level is complicated by the fact that the definition of reading, as it is represented in tests, changes over the school years (Curtis, 1980; Mason & Dunning, 1986). A reading readiness test intended for kindergarten and entering first-grade children may assess ability to identify a few sight words, and name letters and identify their sounds. The reading test to be used at the end of the first grade contains words that children are almost sure to know, if they are able to identify them, and sentences in simple syntax about matters familiar to most first graders. Tests of reading to be used in later grades contain increasingly difficult vocabulary, less likely to be known even when the words can be pronounced, expressed in more complex syntax, about matters less likely to be familiar from firsthand experience, and more often demanding reasoning beyond the given.

The complication introduced by the shifting definition of reading is that one could, for instance, have in kindergarten a nearly perfect measure of letter-naming ability and yet have an imperfect assessment of the talents that will be entailed later by a first-grade reading achievement test. This point was demonstrated by Mason and Dunning (1986), who traced children's reading growth over two years. Two distinct components of reading emerged by the end of first grade, one related to letter and word recognition, and the other to reading and listening comprehension. They loaded as separate factors in a factor analysis and were predicted by different tasks. Letter and word recognition was predicted by letter and sight word tests given in kindergarten and at the beginning of first grade, while comprehension was predicted by oral language and vocabulary knowledge measured at the beginning of kindergarten.

Provided certain assumptions can be met [see below], aggregating data at the classroom level is an approach which reduces problems of fallible measurement. But, it does not eliminate them entirely. Moreover, as a general rule, analyses involving classroom aggregates do not permit inferences about the magnitude, or even the direction, of effects on the performance of individual students, and conversely analyses of individual level data do not permit inferences about class level effects (Burstein, 1980; Pedhazur, 1982). That this is so was graphically illustrated in a "multilevel" reanalysis of the data from the Beginning Teacher Evaluation Study (Marliave, Fisher, & Dishaw, 1978). The conclusion usually cited from this study--that a low error rate is positively related to achievement gains--comes from individual level analyses. However, when the data were analyzed at the class level, a low error rate had significant negative relationships to achievement gains in reading. This pattern of findings fits well with our hypothesis (see also Burstein, 1980).

Confusing the picture, though, is the fact that a multilevel reanalysis of the data from the First-grade Reading Group Study by Martin, Veldman, and Anderson (1980) showed significant positive relationships between a low error rate and reading gains, whether the unit of analysis was the whole class, the reading group, or the individual child. How are these findings to be reconciled with those that came from the Beginning Teacher Evaluation Study reanalysis? To begin with, as scholars such as Brophy (1983), Burstein (1980), Fisher et al. (1978), and Weinstein (1976) have noted, no completely confident interpretation of data from intact, natural groups is ever possible.

With this caveat clearly in mind, we nevertheless offer the following interpretive rule of thumb: To the extent that the classes [for other units of aggregation] participating in a study can be considered

to be essentially random samples from the same population, then between-class variation can be judged to be relatively free from the influence of imperfectly measured differences in initial ability and other background characteristics. On the other hand, to the extent that there have been constraints on assignment of children to the classes included in a study, then between-class variation could very well reflect background differences among the children, and this ought to be the default assumption.

Examples of constraints on assignment of children to classes that are clearly related to ability to learn to read are pooling classes from schools that use tracking systems, in which children are assigned to classes based on judgements of ability, or pooling classes from schools serving neighborhoods with different proportions of low-income and middle-income families. If it were the only constraint, the latter problem might be solved by removing between-school variation, but we can't conceive of how to cope with the first one.

Applying the rule of thumb, the one certain conclusion we have been able to reach is that the reading group is a unit of aggregation in which background differences in ability and differences in treatment are hopelessly intertwined. We will not attempt to reconcile the Beginning Teacher Evaluation Study and the First-grade Reading Group Study. This should be done following more analysis by people who know the studies better than we do.

In the meantime, to recapitulate, we began with the simple assumption that error rate on classroom reading tasks reflects children's level of reading ability. Since one-shot tests are always imperfect indicators of initial reading level, we hypothesized that classroom error rate adds to the prediction of later achievement because it provides further information about ability not contained in previous test scores. This hypothesis offers a parsimonious explanation for the negative correlation between error rate and achievement, after initial reading level has been statistically "controlled"; it resolves the apparent paradox that both a low error rate and a fast pace are associated with increased achievement; it explains subtle features of established findings, for example, that the negative correlation between error rate and achievement is highest at the first grade.

What is the true relationship between errors and growth in reading? If our hypothesis is correct, the relationship is certainly less negative than face interpretation of existing results would suggest. Quite possibly, as we will explain in the next section, the true relationship between errors and growth is positive under some conditions. If our hypothesis is correct, a number of findings in the literature on classroom learning and instruction may have to be reinterpreted. Notice, though, that the hypothesis could be true but not very important. Whether it identifies a minor factor or a major one remains to be determined. Notice, too, that the idea that classroom error rate provides additional information about reading level, not readily available in initial measures, is not inconsistent with the ideas that error rate also reflects in part the quality of the teaching, the character of group processes, and the nature of materials and tasks.

Could Errors Actually Help Students Learn?

Easy acceptance of the belief that errors inhibit learning may reflect a vestigial influence of behaviorism, which waxed strongly in education during the 1950s and 60s. Although the behaviorist world view generally waned during the 70s, it seems to have had an enduring attraction among teacher effectiveness researchers who have done process-product studies in reading, perhaps because of the good showing of programs such as DISTAR (Stallings & Kaskowitz, 1974).

The major educational innovation issuing from behaviorism was programmed instruction. One of the first principles of programmed instruction was to keep the error rate low (Markle, 1964), although research on programmed instruction never succeeded in demonstrating that a low error rate was important. Ironically, what the research consistently showed, instead, was that features of programs intended to minimize errors often inhibited learning by allowing students to short circuit the thinking

required to comprehend and master material (e.g., Anderson, 1970; Anderson, Kulhavy, & Andre, 1972; Holland, 1967).

Throughout the period when behaviorism held sway, there were occasional demonstrations that people learn from their mistakes, especially when attempting to understand meaningful subject matter, but the world was not listening. Notably, Guthrie (1971) found increased learning when feedback followed wrong responses but not when feedback followed correct responses, concluding that the principle of minimizing error rate required rethinking.

Cognitive science provides conceptual tools far beyond those of behaviorism with which to understand errors. The concepts that we shall invoke are *tension* and *resolution*, which harken back to Gestalt psychology (see Koffka, 1935; Kohler, 1947), but which also have been given expression by contemporary cognitive scientists (see especially, Auble, Franks, & Soraci, 1979). The basic idea can be diagrammed as follows:

TENSION--->RESOLUTION--->LEARNING

Tension increases attention and instigates deeper processing or a greater "effort after meaning" (Bartlett, 1932). To state the idea in simple terms, tension leads students to "tune in" to instruction. Tension during reading instruction may come from many sources, including suspense the author has built into a story or the demand for a public performance. Our conjecture is that oral reading errors and mistakes in answering questions can be a source of creative tension. Tension during a reading lesson may be resolved by thinking the student does following an error, by further reading of the text, by the contribution of a classmate, or by feedback and explanation offered by the teacher.

Within limits that we do not claim to know, the greater the tension, the greater is the potential for learning. Even a small amount of tension produces measurable facilitation in learning. Anderson, Goldberg, and Hidde (1971; see also, Kane & Anderson, 1978) had college undergraduates read sentences in which the last word was strongly determined by the rest of the sentence, such as *The monkey peeled the yellow banana*. Under some conditions, there was a blank in the place of the last word and students supplied the word as they read. Usually the incomplete sentences were read as smoothly as the complete ones. In several experiments that varied materials and procedures, students consistently performed better on cued recall tests when they had to fill blanks in sentences instead of reading already completed sentences. The blank created a small tension, in this case a tension that was easily resolved, but one which nevertheless, when resolved, increased learning.

In similar experiments, Auble et al. (1979; also Auble & Franks, 1978) had undergraduates read a number of sentences, such as *The party stalled because the wire straightened*, and a few seconds later read an additional clue such as *corkscrew*. The students who got this arrangement performed considerably better on a later test than students who read *The party stalled because the corkscrew wire straightened* for the same total amount of time. Auble et al. attributed the enhanced learning when *corkscrew* was not in the sentence to an "aha" experience, which depended upon being in a state of puzzlement that resolved into a state of comprehension.

How do errors figure into the equation? One angle on this question was investigated by Kane and Anderson (1978), who composed two kinds of sentences, "determined" and "undetermined." Again, determined sentences, such as *The physician noted the time on his wrist watch*, were so called because the last word was determined by the rest of the sentence. Matched with each determined sentence was an undetermined sentence, in this case, *The physician asked the patient if he had a watch*, so called because it was impossible to get the last word from the rest of the sentence. Some students saw determined sentences, some saw undetermined sentences, and a third group got a mixture consisting of both determined and undetermined sentences. Half of the students completed sentences from which the last word was missing. They were required to give an overt answer; a few seconds later the correct word was displayed. The other half of the students read whole sentences.

Table 1 summarizes results from the Kane and Anderson (1978) study. As in the Anderson et al. (1971) experiments, filling blanks in determined sentences, a task on which the students almost never made a mistake, strongly facilitated learning. The shocking new finding was that filling blanks in undetermined sentences, a task on which students were almost never correct, was equally facilitative. In a second experiment, the students were tested 10-14 days after encountering the sentences. Those who had supplied the last words of undetermined sentences still did significantly better than those who had only read undetermined sentences, but there now was an advantage for having supplied words in determined sentences.

[Insert Table 1 about here.]

That those who struggled with undetermined sentences did as well on immediate tests as those who got determined sentences is surprising because the undetermined sentences were less coherent. Consider the following two sentences:

The bald farmer put the crackers on the top shelf.

The tall farmer put the crackers on the top shelf.

The second sentence is more coherent because it affords an additional link between the agent and the location. Most readers will make the connection that the tallness of the farmer enables his reaching the top shelf. The additional link may be thought of as relating otherwise arbitrary information.

Coherence is another concept that can be traced to Gestalt psychology, and it is alive today in such ideas as *semantic integration* and *precise elaboration*. Theory and data converge on the conclusion that coherent information is more learnable and memorable (see, for instance, Anderson & Pearson, 1984; Anderson & Reder, 1979; Reder & Anderson, 1980; Stein & Bransford, 1979).

There are probably trade-offs between creating tension and establishing coherence. Tension is created when materials, tasks, and problems are challenging. But challenge brings with it the risk of errors, and errors can interfere with the student's constructing maximally coherent representations. Nonetheless, under some conditions the net effect of errors could well be positive.

Microanalysis Of Errors During Reading Lessons

The only real hope of determining whether errors have a causal relationship to reading comprehension is in *microanalyses* of lessons. As the word suggests, a microanalysis gives a fine-grained picture of events. An additional important characteristic, according to Au and Mason (1981), is that a microanalysis of lessons involves *proximal indices* of student performance--that is, measures of performance that are intimately connected to teaching behavior, aspects of group processes, or characteristics of tasks and materials that presumably influence performance. According to this criterion, the Anderson et al. (1979) study, for instance, does not qualify as a microanalysis, even thought it involved a rather comprehensive and detailed analysis of classroom transactions. It does not qualify because the measure of performance was end-of-year achievement instead of one or more proximal indices.

We have recently completed a microanalysis of small-group reading lessons. This study was part of a larger program of research, the long-term goal of which is to determine the aspects of students' ability, teachers' behavior, group processes, and materials that converge at a moment to determine whether particular information will be comprehended, learned, and later remembered. This paper contains the first, preliminary report of our most recent study. We will summarize only the data that pertain to oral reading errors.

Participating in the study were 149 children in six third-grade classes [including one combination second-third grade class and one combination third-fourth grade class] from two schools in a small midwestern city. One school serves a middle-income neighborhood, the other a low-income neighborhood. About 75% of the children were white and about 25% were black, with a few other minorities. According to a standardized test, the sample was above the national average in reading level but showed a wide spread of ability.

As part of the study, the children received four reading lessons in their regular reading groups. Each lesson covered a grade-appropriate story that was not in the basal reader used in the school. During two lessons, the children answered a prediction question before reading each page of the text. The question was always about a matter important to the plot and was always answered, directly or indirectly, on the upcoming page. During the other two lessons, the stress was on accurate oral reading. The questions asked between pages were about the properties of individual words--phonics, word structure, and the like--and could be answered on the basis of the page that had just been read. For shorthand, these teaching approaches will be called *meaning emphasis* and *surface emphasis*, respectively. The teaching approach was "counterbalanced" with lesson order and stories. In other words, half the classes got meaning emphasis lessons with the first two stories and surface emphasis lessons with the second two stories, while the remaining classes got surface emphasis lessons with the first two stories and meaning emphasis lessons with the second two. Thus, at its core, this was a true experiment.

During all lessons, children took strictly ordered turns reading aloud pages from the day's story. When there was a meaning emphasis, oral reading errors were ignored, unless meaning was disrupted, in which case the teacher supplied the word and the child went right on reading. When there was a surface emphasis, all reading errors were corrected by the teacher and the child repeated the correct word before continuing.

The lessons were taught by the regular classroom teacher based on lesson guides prepared by the research team. The teacher recorded total lesson time and, for each page, the name of the child who was taking the "active" turn--the child reading aloud and responding to questions and feedback--and the oral reading errors made by that child. Other measures included group and individual comprehension level and word reading fluency, various story and page properties, and several indices of story comprehension and word learning.

[Insert Figure 1 about here.]

Presented in Figure 1 is a partial *causal model* to account for the approximately 3,000 observations recorded during surface emphasis lessons. An observation is defined as the performance of an individual child on a single page of text. Summarized is the performance of "nonturntakers"--children in the group who were reading silently and following along. The arrows mark "causal paths." That is, an arrow indicates that the first factor is hypothesized to be the cause or part of the cause of the second factor. The numbers beside the arrows are coefficients that show the strength of the relationships as determined in the study. These are unstandardized regression coefficients [standard errors in parentheses] where the ability measures are represented in stanines, oral reading errors in logits, and comprehension in radians. The comprehension measure is the proportion of information on the page that could be recalled after the lesson. The variation due to story and page properties has been removed.

Based on the reasoning sketched earlier, we had anticipated that oral reading errors would improve comprehension when the lesson had a surface emphasis. If the goal is saying the words right, then a failure to meet this norm produces tension, and resolution of the tension enhances comprehension. Figure 1 shows that this prediction was confirmed: The coefficient representing the connection between oral reading errors and comprehension was +.10 [.03], a highly significant relationship.

The finding that nonturntakers benefited from reading errors held only when the lesson had a surface emphasis. When the lesson had a meaning emphasis, the coefficient representing the relationship between oral reading errors and comprehension for nonturntakers was close to zero, -.01 [.04]. This result was also expected on the basis of theory and findings from a previous study (Anderson, Mason, & Shirey, 1984). The explanation is that, when there is an emphasis on the meaning of a story, oral reading errors are irrelevant to the task and, therefore, they do not create much tension and do not affect comprehension. This is provided, of course, that the teacher corrects the errors that would interfere with the development of a coherent text representation.

What about the influence of errors on the comprehension of the active turntakers, the children who themselves were making the errors? Here, the relationship was uniformly negative. The coefficients were -.08 [.08] and -.10 [.10] for the surface and meaning emphasis lessons, respectively. It is important to stress that this negative effect was more than countered by the significant, general benefit from being the active turntaker. In other words, the net effect of an active turn was positive, even when the child made some mistakes. The net influence was negative only when the child really butchered the passage.

How is the negative influence of active turntakers' errors on their own comprehension to be understood? There are several possibilities. The first is the one developed at length in the preceding sections: Oral reading errors are an additional indicator of ability, and poor readers do not comprehend as well as good readers. A second possibility is emotional: Children making a lot of mistakes get flustered and their attention is drawn away from meaning. A third possibility is cognitive: Despite the teacher's feedback, errors interfere with the development of a coherent understanding of the story. The study contains no clues that would permit a choice among these possible explanations.

The findings from the study that has just been summarized are entirely consistent with the findings from an earlier study by our research group (Anderson et al., 1984), in which 264 third graders in 12 classes from five elementary schools in another small midwestern city participated. The children received a single lesson that had either a meaning emphasis or a surface emphasis. The materials were sentences that varied from the first-grade level to the seventh-grade level on the Fry readability scale. The measure of learning was recall of the sentences after the lesson. When the lesson had a surface emphasis, but not when it had a meaning emphasis, the higher the readability grade level of the sentences the greater was the recall. This must mean that reading errors enhanced learning. Similarly, when the lesson had a surface emphasis, but not when it had a meaning emphasis, the less fluent the active turntaker the greater was the recall of the sentences. Again, this must mean that errors enhanced learning, provided the task emphasized fluent, accurate reading. As in our most recent study reported earlier, being the active turntaker significantly benefited learning.

To summarize, evidence from two microanalytic studies suggests that oral reading errors during a reading lesson contribute to comprehension among nonturntakers when the task is accurate, fluent reading. In contrast, oral reading errors do not seem to affect the comprehension of nonturntakers when the task is story understanding. Both of these results are readily interpretable in terms of a theory that invokes the concepts of tension and resolution. Oral reading errors are negatively related to comprehension for active turntakers, a fact for which there are several possible interpretations. However, no matter which interpretation is more correct, there is a net positive benefit to comprehension from being the active turntaker for all but children who make far above average numbers of errors.

Toward a Richer Concept of Success and Failure

All errors may not be functionally equivalent. An oral reading error may not be equivalent to an error on a skill sheet assignment or an error in response to a question during a class discussion. The tension-resolution hypothesis assumes errors that result from effortful application of strategies on the part of the reader, and errors resulting from unmotivated, haphazard guesswork may have different consequences for learning (J. E. Brophy, personal communication, November 1986). A child's history of success or failure in reading may be important. Occasional errors within the context of an otherwise smooth performance may present an altogether different circumstance from that in which a child stumbles over many words, struggling to maintain coherence. The true relationship between errors and growth in reading may indeed be curvilinear; beyond some level of error, additional errrors may cause more confusion and frustration than learning.

E. H. Hiebert (personal communication, January 1987) has developed a framework for analyzing issues of success and failure on classroom tasks such as the foregoing. The framework involves *learner* variables, *task* variables, and *context* variables. Learner variables include the ability, self-esteem, maturity, and history of success and failure of the students. Task variables include goals and purposes set by the teacher, the degree of problem solving required, and how interesting students find the task and material. Context variables include how turns for reading or answering questions are determined, whether at a particular moment a certain student is the turntaker or a nonturntaker, whether students' responses are public or private, and the culture of the reading group.

The two microanalytic studies summarized in the preceding section included a task variable--whether the teacher maintained an emphasis on understanding the material or an emphasis on accurate, fluent oral reading. As already recounted, this variable strongly conditioned the influence of oral reading errors. However, another task variable included in one of the studies (Anderson et al., 1984)--the rated interest of the reading material--had no effect on the influence of oral reading errors. These two studies also included a context variable. The times children were turntakers were compared with the times they were nonturntakers. This variable, too, conditioned the influence of oral reading errors. Other studies have examined error rate in the light of task and context variables, but the major point to be made is that systematic research on these variables has not progressed very far.

More attention has been paid to learner variables. Educators who work with students who could be described as immature, low in ability or self-esteem, or who have a history of poor school performance are very receptive to the message from instructional theory and research that a low error rate is especially important for children in these categories, because they report that it squares with their experience (I. W. Gaskins, personal communication, February 1987). The systematic research basis for this conclusion is thin, however. The best available evidence comes from a further analysis of the Beginning Teacher Evaluation Study reported by Marliave and Filby (1985): Fall to spring gains in reading achievement were strongly related to a high rate of success on classroom tasks for low-ability second graders, less strongly related for average- and high-ability second graders or fifth graders of any ability level. We are dubious about this evidence because we believe the analysis upon which it is based is particularly vulnerable to the criticism presented earlier in this paper, namely, that a high success rate provides additional information about the reading level of the children. However, it must be noted that Marliave and Filby (1985, pp. 226-227) maintain the opposite is true.

In any event, no evidence was reported from the First-grade Reading Group Study by Anderson et al. (1979) that children with low reading readiness are especially sensitive to error rate. Hoffman et al. (1984) found that low-ability second graders in reading groups experiencing a high rate of oral reading errors were less likely than high-ability children to pause spontaneously following an error and less likely to be given the opportunity to make self-corrections by their teacher. However, Hoffman et al. did not report that the gains low-ability children made on a reading achievement test

were especially sensitive to rate of oral reading errors. Nor did we find in the microanalytic studies summarized in this paper that the amount low-ability children learned from a lesson depended any more or less on rate of oral reading errors than did the amount high-ability children learned.

To summarize, different kinds of errors probably have different consequences for progress in reading, and these consequences probably depend upon learner, task, and context variables.

Conclusion

One thesis of this paper is that the association between a low rate of errors on classroom tasks, especially a low rate of oral reading errors, and gains in reading achievement may be an epiphenomenon. It may be no more than a reflection of the fact that good readers usually make fewer mistakes than poor readers on any task that involves reading. Of course, investigators have tried to factor out initial reading level, but--without any exception of which we are aware--they have failed to take adequate account of the fact that measures of reading are always imperfect. All classroom error rate may do is provide additional and more reliable information about children's reading ability. Therefore, the grounds for the almost taken-for-granted conclusion that a low error rate is on the causal path to increased reading achievement are actually quite insubstantial.

A creditable case can be made that, within unknown limits, oral reading errors--and maybe errors on other types of classroom tasks--ought to *help* rather than hurt growth in reading. An oral reading error followed by feedback appears to fit the pattern of tension followed by resolution that other research has shown improves comprehension, learning, and memory.

Microexperimental analyses of third-grade reading lessons confirm that oral reading errors do have a significant, positive influence on the comprehension of nonturntakers. As anticipated, errors improved comprehension when the task was accurate oral reading but made no difference when the task emphasized story understanding. It would be premature, though, to leap from these results to recommendations for practice, because these studies have their limitations. Notably, the mean rate of oral reading errors was very low in our most recent study, only 1.4% [range from 0% to 16.4%]. This error rate is close to what reading educators regard as the *independent level*--material easy enough that children can read it on their own without becoming frustrated. So, the low error rate is a serious limitation. Although the error rate was higher in our earlier study (Anderson et al., 1984), research that systematically examines variations in error rate is urgently needed.

One's confidence that the principal findings of the two microanalyses briefly summarized here are generalizable is bolstered by the fact that another microanalysis that examined oral reading errors also obtained positive results. In an exemplary study, Eder and Felmlee (1984; see also Felmlee & Eder, 1983) videotaped 16 lessons in a single first-grade classroom. They discovered that sustained attention during a reading turn was strongly associated with a high rate of error. That is to say, the more oral reading errors the current turntaker was making, the more likely the children were to stay attentive. Notice that these results invite an interpretation favoring the tension-resolution theory that we have advanced.

The tension-resolution hypothesis is not the only explanation for the positive, proximal influence of oral reading errors on comprehension, although we believe it is the one that best fits available data. An alternative explanation is that reading errors may increase students' time on sections of text, provide occasions for teacher feedback, and allow more opportunities for rehearsal (see Anderson et al., 1984). Weighing against this alternative is Felmlee and Eder's (1983) finding that children become more inattentive as the length of a reading turn increases.

If the positions developed in this paper are correct, conclusions about error rate are not the only widely accepted beliefs about classroom instruction that may have to be revised. On the basis of the studies reviewed earlier, sustaining feedback is recommended to teachers while terminal feedback is

proscribed. Yet a close look at the results presented by Hoffman et al. (1984) plainly shows that teachers tend to give sustaining feedback following the kinds of errors characteristic of good readers and terminal feedback following the kinds of errors characteristic of poor readers. The association between type of feedback and end-of-year achievement could turn out to be just another indirect reflection of the student's initial reading level.

Another accepted conclusion that our analysis leads us to question is the idea that early teacher decisions about children become self-fulfilling prophesies. Starting with Weinstein's (1976) important study, the theory has been articulated that when children enter the first grade their abilities as readers are quite plastic. Conventional American practice requires that children be sorted into reading groups. However, the theory continues, the criteria for group placement are insufficient, and decisions may be influenced by such factors as the children's ethnic and social class origins. Once placed in a reading group, powerful forces associated with group membership--differences in pace, quality of teaching, standards of acceptable performance, group attention norms--are theorized to take hold and determine a child's eventual level of success. Variations in the instruction that different reading groups receive have been amply demonstrated (e.g., Allington, 1983; Hiebert, 1983). However, the linchpin assumption of the theory--that placement in groups is underdetermined--bears closer inspection. The assumption rests on the finding of Weinstein and others (see Barr et al., 1983) that group placement is a better predictor of first-grade achievement than scores on a reading readiness test. However, to repeat, a readiness test is only an imperfect indicator of reading level. We contend that error rate on daily classroom reading tasks, which teachers presumably use to compose groups, is a quite satisfactory indicator of current ability and, retrospectively, is probably a better indicator of initial ability than a readiness test.

Because of the way research questions have been framed, the answers for instructional policy that can be gleaned from previous research on classroom error rate are inconclusive at best, and wrong at worst. The proper question is--for a child with a given level of current ability--does an error represent an obstacle to further growth, or an opportunity? Our two studies and the one by Eder and Felmlee (1984) suggest an answer to this question that runs very much against conventional wisdom. These are small studies, to be sure, and in need of replication. More important, these studies leave many questions about errors on classroom tasks begging for dependable answers. All we can claim at the present time is that we have interesting preliminary evidence that suggests that--contrary to popular opinion--oral reading errors may actually contribute to learning.

References

Allington, R. L. (1983). The reading instruction provided readers of differing reading abilities. *Elementary School Journal, 83*, 548-559.

Anderson, L. M., Evertson, C. M., & Brophy, J. E. (1979). An experimental study of effective teaching in first-grade reading groups. *Elementary School Journal, 79*, 193-223.

Anderson, R. C. (1970). Control of student mediating processes during verbal learning and instruction. *Review of Educational Research, 40*, 349-369.

Anderson, R. C., Goldberg, S. R., & Hidde, J. L. (1971). Meaningful processing of sentences. *Journal of Educational Psychology, 62*, 395-399.

Anderson, R. C., Hiebert, E. H., Scott, J. A., & Wilkinson, I. A. G. (1985). *Becoming a nation of readers: The report of the Commission on Reading*. Champaign, IL: Center for the Study of Reading.

Anderson, R. C., Kulhavy, R. W., & Andre, T. (1972). Conditions under which feedback facilitates learning from programmed lessons. *Journal of Educational Psychology, 63*, 186-188.

Anderson, R. C. Mason, J., & Shirey, L. (1984). The reading group: An experimental investigation of a labyrinth. *Reading Research Quarterly, 20*, 6-38.

Anderson, R. C., & Pearson, P. D. (1984). A schema-theoretic view of basic processes in reading comprehension. In P. D. Pearson (Ed.), *Handbook of reading research* (pp. 255-291). New York: Longman.

Anderson, J. R., & Reder, L. M. (1979). An elaborative processing explanation of depth of processing. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory* (pp. 385-403). Hillsdale, NJ: Erlbaum.

Au, K. H., & Mason, J. M. (1981). Social organizational factors in learning to read. *Reading Research Quarterly, 17*, 115-152.

Auble, P. M., & Franks, J. J. (1978). The effects of effort toward comprehension on recall. *Memory and Cognition, 6*, 20-25.

Auble, P. M., Franks, J. J., & Soraci, S. A. (1979). Effort toward comprehension: Elaboration or "aha!"? *Memory and Cognition, 7*, 426-434.

Barr, R., Dreeben, R., & Wiratchai, N. (1983). *How schools work*. Chicago, IL: University of Chicago Press.

Bartlett, F. C. (1932). *Remembering*. Cambridge, MA: Cambridge University Press.

Betts, E. A. (1946). *Foundations of reading instruction*. New York: American Book Company.

Blaxall, J., & Willows, D. M. (1984). Reading ability and text difficulty as influences on second graders' oral reading errors. *Journal of Educational Psychology, 76*, 330-341.

Brophy, J. E. (1983). Research on the self-fulfilling prophecy and teacher expectations. *Journal of Educational Psychology, 75*, 631-661.

Burstein, L. (1980). The analysis of multilevel data in educational research and evaluation. In D. Berliner (Ed.), *Review of research in education* (Vol. 8, pp. 158-233). Washington, DC: American Educational Research Association.

Butler, S. R., Marsh, H. W., Sheppard, M. J., & Sheppard, J. L. (1985). Seven-year longitudinal study of the early prediction of reading achievement. *Journal of Educational Psychology, 77*, 349-361.

Cunningham, J. W. (1985). Three recommendations to improve comprehension teaching. In J. Osborn, P. T. Wilson, & R. C. Anderson (Eds.), *Reading education: Foundations for a literate America* (pp. 255-274). Lexington, MA: Lexington Books.

Curtis, M. (1980). Development of components of reading skill. *Journal of Educational Psychology, 72*, 656-669.

Eder, D., & Felmlee, D. (1984). The development of attention norms in ability groups. In P. Peterson, L. Wilkinson, & N. Hallinan (Eds.), *The social context of instruction* (pp. 189-208). Orlando, FL: Academic Press.

Felmlee, D., & Eder, D. (1983). Contextual effects in the classroom: The impact of ability groups on student attention. *Sociology of Education, 56*, 77-87.

Fisher, C. W., Filby, N. N., Marliave, R., Cahen, L. S., Dishaw, M. M., Moore, J. E., & Berliner, D. C. (1978). *Teaching behaviors, academic learning time, and student achievement: Final report of phase III-B, Beginning Teacher Evaluation Study* (Tech. Rep. V-1). San Francisco: Far West Laboratory for Educational Research and Development.

Guthrie, J. T. (1971). Feedback and sentence learning. *Journal of Verbal Learning and Verbal Behavior, 10*, 23-28.

Hiebert, E. H. (1983). An examination of ability grouping for reading instruction. *Reading Research Quarterly, 18*, 231-255.

Hoffman, J. V., O'Neal, S. F., Kastler, L. A., Clements, R. O., Segel, K. W., & Nash, M. F. (1984). Guided oral reading and miscue focused verbal feedback in second-grade classrooms. *Reading Research Quarterly, 19*, 367-384.

Holland, J. G. (1967). A quantitative measure for programmed instruction. *American Educational Research Journal, 4*, 87-101.

Juel, C., Griffith, P. L., & Gough, P. B. (1986). Acquisition of literacy: A longitudinal study of children in first and second grade. *Journal of Educational Psychology, 78*, 243-255.

Kane, J. H., & Anderson, R. C. (1978). Depth of processing and interference effects in the learning and remembering of sentences. *Journal of Educational Psychology, 70*, 626-635.

Koffka, K. (1935). *Principles of gestalt psychology*. New York: Harcourt, Brace.

Kohler, W. (1947). *Gestalt psychology*. New York: Liveright.

Leinhardt, G., Zigmond, N., & Cooley, W. W. (1981). Reading instruction and its effects. *American Educational Research Journal, 18*, 343-361.

Linn, R. L., & Werts, C. E. (1982). Measurement error in regression. In G. Keren (Ed.), *Statistical and methodological issues in psychology and social sciences research* (pp. 131-154). Hillsdale, NJ: Erlbaum.

Markle, S. M. (1964). *Good frames and bad: A grammar of frame writing*. New York: Wiley.

Marliave, R., & Filby, N. N. (1985). Success rate: A measure of task appropriateness. In C. W. Fisher & D. C. Berliner (Eds.), *Perspectives on instructional time* (pp. 217-236). New York: Longman.

Marliave, R., Fisher, C. W., & Dishaw, M. M. (1978). *Academic learning time and student achievement in the B-C period* (Tech. Note V-2a). San Francisco: Far West Laboratory for Educational Research and Development.

Martin, J., Veldman, D. J., & Anderson, L. M. (1980). Within-class relationships between student achievement and teacher behaviors. *American Educational Research Journal, 17*, 479-490.

Mason, J. M., & Dunning, D. (1986). *Proposing a model to relate preschool home literacy with beginning reading achievement*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Pedhazur, E. J. (1982). *Multiple regression in behavioral research*. New York: Holt, Rinehart and Winston.

Reder, L. M., & Anderson, J. R. (1980). A partial resolution of the paradox of interference: The role of integrating knowledge. *Cognitive Psychology, 12*, 447-472.

Rosenshine, B., & Stevens, R. (1984). Classroom instruction in reading. In P. D. Pearson (Ed.), *Handbook of reading research* (pp. 745-798). New York: Longman.

Stallings, J. A., & Kaskowitz, D. (1974). *Follow-through classroom observation evaluation 1972-1973*. Menlo Park, CA: Stanford Research Institute.

Stein, B. S., & Bransford, J. D. (1979). Constraints on effective elaboration: Effects of precision and subject generation. *Journal of Verbal Learning and Verbal Behavior, 18*, 769-777.

Walberg, H. J., & Tsai, S (1984). Reading achievement and diminishing returns to time. *Journal of Educational Psychology, 76*, 442-451.

Weinstein, R. S. (1976). Reading group membership in first grade: Teacher behaviors and pupil experience over time. *Journal of Educational Psychology, 68*, 103-116.

Wilkinson, I., Wardrop, J. L., & Anderson, R. C. (in press). Silent reading reconsidered: Reinterpreting reading instruction and its effects. *American Educational Research Journal*.

Zeller, R. A., & Carmines, E. G. (1980). Measurement in the social sciences. Cambridge, MA: Cambridge University Press.

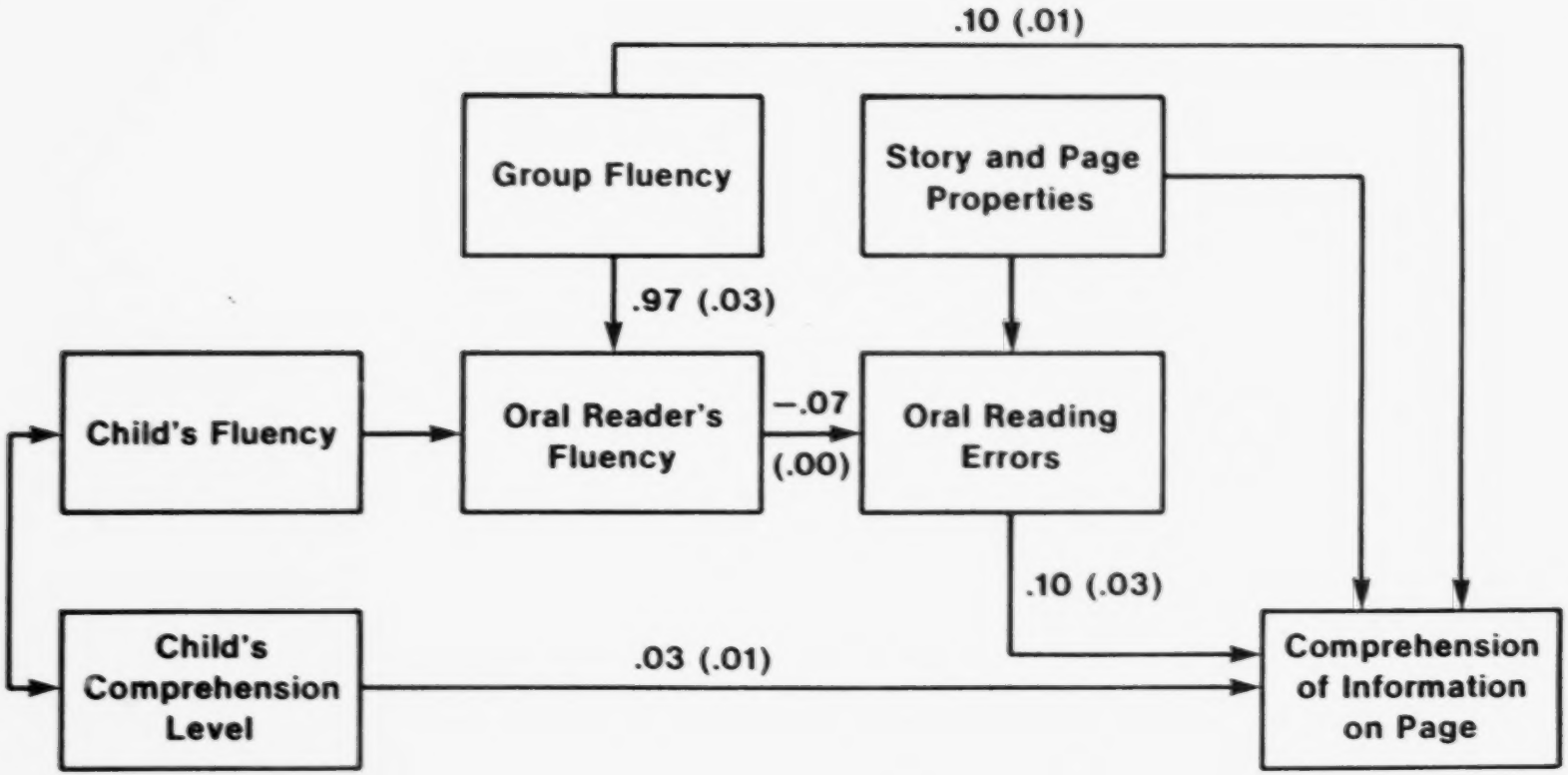
Table 1

Mean Proportions Recalled in First Experiment as a Function of List Type and Experimental Task (from Kane & Anderson, 1978)

	List Type		
Experimental Task	Determined	Undetermined	Mixed
Sentence Completion	.77	.78	.76
Reading Only	.69	.66	.70

Figure Caption

Figure 1. Partial causal model of relations among reading abilities, oral reading errors, and comprehension of nonturntakers during surface emphasis lessons.


.10 (.01)
Group Fluency
Story and Page Properties
.97 (.03)
Child's Fluency
Oral Reader's Fluency
−.07
(.00)
Oral Reading Errors
.10 (.03)
Child's Comprehension Level
.03 (.01)
Comprehension of Information on Page